



# Impact of STIR's programming on teacher motivation and student learning

Endline report

July 2018<sup>1</sup>

---

<sup>1</sup> This report has been prepared by the IDinsight team. Please direct all correspondence regarding this report to [heather.lanthorn@idinsight.org](mailto:heather.lanthorn@idinsight.org).

## Executive summary

### STIR Program and Evaluation Design

We report results from randomized evaluations of STIR’s programming in Delhi and Uttar Pradesh (U.P.). STIR seeks to improve teacher motivation and classroom practice by organizing teachers into local networks. These networks hold monthly, guided meetings where teachers discuss principles of classroom practice and share ideas for how to improve their teaching. In Delhi, STIR worked with teachers at private schools with monthly fees less than \$17 (sometimes referred to as “affordable private schools”) and STIR staff directly organized and guided the monthly network meetings. In U.P., STIR worked with government schools and trained and coached volunteer government school teachers to organize and guide the meetings. In both Delhi and U.P., schools included grades from 1<sup>st</sup> to 8<sup>th</sup> standard and roughly 20% of teachers participated in STIR meetings.<sup>2</sup>

We randomized the offer of STIR programming in two stages. First, schools were randomly assigned to either treatment or control. We then grouped nearby treatment schools into clusters and randomly assigned each cluster of schools to receive either the STIR “standard” model or the STIR “exploratory” model. In addition to the network meetings, teachers in the exploratory model also received non-financial incentives such as recognition from local officials. We collected data on classroom practices, teacher motivation, and student learning outcomes at baseline, one year later, and two years later.

All findings reported below are school-level results. That is, we compare *all* teachers and students in treatment schools, regardless of whether the teacher participated in STIR meetings, to *all* teachers and students in control schools. In the body of the report, we also present estimates of the effect of STIR on teachers who participate in STIR meetings.

### Findings from Delhi Private Schools

**In Delhi, we find that the offer of STIR programming led to improved math learning outcomes.** Students in STIR schools (standard + exploratory combined) increased math learning levels by .1 standard deviations (p-value: 0.02) and students in the standard treatment arm increased math levels by .15 standard deviations (p-value: 0.04) compared to students in control schools. These effects appear to be driven mainly by poor performers achieving a basic math learning level. **We find no effect on Hindi learning outcomes in Delhi.**

**In Delhi, we also find suggestive evidence that STIR led to increased teacher motivation.** STIR led to a 0.13 standard deviation increase in an overall index measuring teacher motivation among teachers in the standard treatment arm (p-value: <0.01). In addition, we find effects on a sub-index which sought to measure “growth mindset,” one of three analyzed sub-indices. STIR led to a 0.15 standard deviation increase on the

---

<sup>2</sup> According to STIR, participation was limited to roughly 30% of teachers at each treatment school. Our understanding is that, in practice, this cap was rarely binding.

growth mindset sub-index among teachers in STIR schools and a 0.18 standard deviation increase on this sub-index among teachers in the standard treatment arm. We do not find significant effects for the overall index for STIR schools or for the two other sub-indices (teacher efficacy and positive professional outlook).

## Findings from U.P. Government Schools

**In U.P. government schools, we find weak evidence of gains in the amount of time teachers spend teaching.** STIR led to a 4 percentage-point increase (p-value: 0.08) in observed teaching time among teachers in STIR schools. In the standard treatment arm, STIR led to an 8-percentage-point increase (p-value: 0.09) in observed teaching time. We characterize this evidence as weak given the large number of outcomes we test for and the relatively large p-values of the results.

**In U.P, we do not find statistically significant effects on teacher motivation, student learning outcomes, or other classroom practices.**

## Conclusions

Our results show that STIR's approach can work but that its effectiveness depends on context, where context may include geography; education systems, financing, and staffing; and program components and approaches to delivery. In Delhi, STIR caused a 0.1 standard increase in math learning outcomes. This result is similar in size to effect sizes from other teacher training and incentive interventions in low- and middle-income countries (McEwan 2015; Snilstveit et al. 2015). In U.P., we find weak evidence that STIR may have increased teaching time and no effect on learning outcomes and several other measures. Unfortunately, we are unable to pin down the source of this difference. There are several large differences in both the context and implementation model between the Delhi and U.P. versions of the program. Our evaluation is unable to disentangle the importance of these differences.

## Limitations

This study has three key technical limitations. First, we experienced a high level of teacher and student attrition. We do not detect differential attrition on observables between treatment and control but cannot rule out differential attrition on unobservable characteristics. Second, we analyze many hypotheses which raises the risk of false positive findings. We correct for multiple hypothesis testing within outcomes families with more than four outcomes but do not correct across outcome families. Third, our classroom observations may be subject to observer effects, as some of the child-friendly measures were explicitly highlighted as part of discussions in the community of practices.